

WestminsterResearch

<http://www.westminster.ac.uk/westminsterresearch>

**An Adaptive Fuzzy Logic System for the Compensation of
Nonlinear Distortion in Wireless Power Amplifiers
Kodogiannis, V., Vaskovic, M. and Budimir, D.**

This is an author's accepted manuscript of an article published in Neural Computing and Applications, doi: 10.1007/s00521-017-2849-3, 2017.

The final publication is available at Springer via:

<https://dx.doi.org/10.1007/s00521-017-2849-3>

The WestminsterResearch online digital archive at the University of Westminster aims to make the research output of the University available to a wider audience. Copyright and Moral Rights remain with the authors and/or copyright owners.

Whilst further distribution of specific materials from within this archive is forbidden, you may freely distribute the URL of WestminsterResearch: (<http://westminsterresearch.wmin.ac.uk/>).

In case of abuse or copyright appearing without permission e-mail repository@westminster.ac.uk

An Adaptive Fuzzy Logic System for the Compensation of Nonlinear Distortion in Wireless Power Amplifiers

Mina Vaskovic¹, Vassilis S. Kodogiannis^{2*} and Djuradj Budimir³

¹ Faculty of Science and Technology, University of Westminster, London W1W 6UW, United Kingdom
E-mail: mina.vaskovic@gmail.com

^{2*} Faculty of Science and Technology, University of Westminster, London W1W 6UW, United Kingdom
E-mail: V.Kodogiannis@westminster.ac.uk, Tel: +44-7775708976
(Corresponding Author)

³ Faculty of Science and Technology, University of Westminster, London W1W 6UW, United Kingdom
E-mail: D.Budimir@westminster.ac.uk

Abstract

Computational intelligent systems are becoming an increasingly attractive solution for power amplifier (PA) behavioural modelling, due to their excellent approximation capability. This paper utilizes an adaptive fuzzy logic system (AFLS) for the modelling of the highly nonlinear MIMIX CFH2162-P3 PA. Moreover, PA's inverse model based also on AFLS has been developed in order to act as a pre-distorter unit. Driving an LTE 1.4 MHz 64 QAM signal at 880 MHz as centre frequency at PA's input, very good modelling performance was achieved, for both PA's forward and inverse dynamics. A comparative study of AFLS and neural networks (NN) has been carried out to establish AFLS as an effective, robust, and easy-to-implement baseband model, which is suitable for inverse modelling of PAs and capable to be used as an effective digital pre-distorter. Pre-distortion system based on AFLS, achieved distortion suppression of 84.2%, compared to the 48.4% gained using the NN-based equivalent scheme.

Keywords: *Linearization; Pre-distortion; RF power amplifiers; Neuro-Fuzzy systems; Neural Networks*

1. Introduction

Modern emerging mobile systems such as High Speed Packet Access (HSPA), Long-Term Evolution (LTE), and Mobile Interoperability for Microwave Access (Mobile WiMAX), require radio transceivers able to support high data rates and throughput. One of the main trends in designing wireless transmitters is the delivery of enhanced transmitter functionalities through digital signal processing (DSP) schemes. Since the modulated signals of these wireless systems have higher peak-to-average ratio (PAR) and broad bandwidth, power amplifiers (PA) with high linearity are thus required [1]. However, the assumption of linearity for wireless communication systems limits their efficiency as such assumption is not realistic due to the nonlinear properties of the front-end components and semiconductor devices in the PA [2]. The simplest method is to back-off the power level to achieve high linearity, but with very low efficiency. In order to increase efficiency and minimize the distortion, many different linearization techniques, such as the feed-forward method [3], feedback method [4], analog predistortion [5] and digital predistortion schemes [6] have been developed. Among these techniques, digital predistortion (DPD) has become the most potential method due to its high accuracy and flexibility, as well as to advances in digital signal processors and digital-to-analog converters. The behavioural modelling and the inverse modelling of PAs are the most

important parts in a DPD system. Various schemes of memory or memory-less models have been reported in literature for behavioural modelling, such as the Volterra series model [7], Look-Up Tables (LUT) scheme [8], Wiener model [9], polynomial model [10], neural network (NN) model [11] and neuro-fuzzy (NF) model [12].

The PA exhibits two nonlinear distortions, which involve the amplitude and phase distortion respectively. Earlier approaches attempted to model the PA behaviour by utilizing two uncoupled NN-based models that attempt to capture amplitude and phase responses separately [13]. An adaptive neuro-fuzzy inference system (ANFIS) for DPD has been discussed in [14], however the adopted scheme was static and no memory effects of PAs were considered. The linearization of PAs with memory effects using ANFIS has been considered in [15]. In that paper, two ANFIS schemes have been employed as pre-distorters for correcting the amplitude and phase distortions separately. Neuro-fuzzy based identification is an effective method for modelling nonlinear systems because it combines the advantages of neural systems and fuzzy logic systems. It is represented by fuzzy IF-THEN rules, while the parameters involved can be updated by a suitable learning algorithm. However, ANFIS's main limitations are the exponential growth of rules subjected to the number of input variables as well as its applicability only to single output problems. The majority of existing neuro-fuzzy schemes, like ANFIS, follow the classic Takagi–Sugeno–Kang (TSK) structure, where only one output is enabled [16].

Currently, many researchers have utilized NNs' multiple input multiple output (MIMO) modeling capabilities, and thus PA modeling has been implemented through the use of single NN MIMO models [17-18]. The advantage of a MIMO network against a two-single structures scheme is also related to the possible adaptation needed in case PA's behavior is modified. The main weakness of the two-single structures scheme is the asynchronous convergence of the two models, where both models do not converge to an optimal model at same time, leading to over- or under-training of one model [17].

To overcome the limitations of ANFIS scheme, the main objective of this paper is to associate the modelling of forward and inverse characteristics of a PA through a MIMO neuro-fuzzy scheme. The proposed "Adaptive Fuzzy Logic System" (AFLS) model, in addition to its MIMO characteristics, includes an alternative defuzzification scheme, and address adequately the "*curse of dimensionality*" rules problem, which is an inherent issue in TSK-based neuro-fuzzy structures. The AFLS scheme is utilized for the modelling and pre-distortion of the MIMIX CFH2162-P3 power amplifier. Modelling procedures have been carried out for the LTE 1.4 MHz 64 QAM signal at 880 MHz as centre frequency, at the 1dB compression point of the device under test (DUT). Experimental results from AFLS are then compared only with similar MIMO-based structures, such as the time-delayed neural network (TDNN) model. Such comparison is considered as an essential practice, as we have to emphasize the need of induction to the area of wireless communication systems, advanced learning-based MIMO modelling schemes, which may have a significant potential for an efficient DPD system.

2. Adaptive Fuzzy Logic System

A fuzzy system approximates an unknown mapping by inference from a set of humanly understandable statements or rules. Fuzzy systems represent the imprecision found in real-world problems using IF-THEN rules expressed in a natural language. Fuzzy set theory is a mathematical tool for translating abstract concepts found in natural language

into computable entities. Such entities are called fuzzy sets. Mathematically, a fuzzy set A is represented by a membership function defined on a domain X given by $\mu_A : X \rightarrow [0,1]$, where A is the fuzzy label or linguistic (value) term describing the variable x and $\mu_A(x)$ represents the grade of membership of x belonging to the fuzzy set A . A fuzzy system generally consists of three parts [19]:

- **Fuzzifier:** used at the front of the system to convert crisp data to fuzzy sets. Fuzzification is the process performed by the fuzzifier. The fuzzification process smooths out the model response, making it less sensitive to specific input values and hence to noise in the input.
- **Fuzzy rule base and Fuzzy inference engine:** fuzzy rule base represents the knowledge base of the system and the fuzzy inference engine acts like a decision making logic system that combines the rules in the rule base according to approximate reasoning theory to produce a mapping from fuzzy sets in the input space to fuzzy sets in the output space.
- **Defuzzifier:** used at the output of the system to convert fuzzy sets into crisp values. The defuzzification stage is needed to obtain a crisp output from the fuzzy output resulting from the inference of rules.

Although NNs have good learning abilities, fuzzy logic systems lack such characteristics. The idea in NF systems is to merge the capabilities of model-free and trainable systems and noise tolerance of neural networks with the ability of dealing with imprecise situations from the fuzzy set theory.

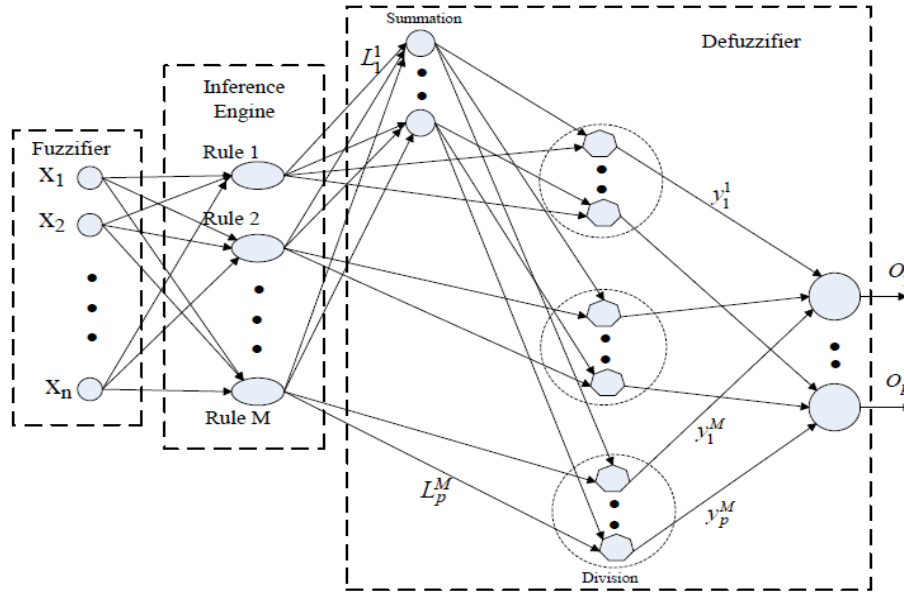


Fig. 1: AFLS Architecture

The majority of NF systems follow a fuzzy rule-table approach that utilise the “look-up table” concept. In those models, an input space is divided into $K_1 \times K_2 \times \dots \times K_n$ fuzzy subspaces, where $K_i, i = 1, 2, \dots, n$ is the number of fuzzy subsets for the i^{th} input variable [20]. As one fuzzy rule is assigned for each of these subspaces, their main drawback is that the number of fuzzy rules increases exponentially with respect to the number of inputs n . ANFIS is a

classical example of such approach, where the number of fuzzy rules is related to the number of input variables as well as the number of membership functions for each input.

Several methods have been used in defuzzification section, such as “centre average (CA)”, “centroid of area (COA)”, centre of sums (COS), *etc.* The COA approach is an optimal answer to defuzzification because it uses all available information to compute the output. One major problem, however, with this approach is its intensive computation and its limited applicability in neurofuzzy systems, due to the presence of a learning algorithm for the tuning of consequent parameters [21]. The CA defuzzifier, on the other hand, is more efficient in terms of implementation, but is only valid for symmetrical output membership functions. Its main disadvantage is that it suffers from not utilising the entire shape of the consequent membership function. The output of a CA defuzzifier is still the same, regardless of whether the shape is narrow or wide [22]. However, like COA, CA is suitable for MIMO fuzzy logic systems. Finally, defuzzification at TSK models follows a different strategy by replacing the fuzzy consequent part with a polynomial function of the input variables. TSK fuzzy model is by far the most popular candidate for single output modelling problems [16]. The proposed MIMO AFLS scheme does not follow TSK’s architecture, as the number of memberships for each input variable is directly associated to the number of rules, hence, the “*curse of dimensionality*” problem is significantly reduced. In addition, an alternative to CA defuzzification approach has been adopted, the area of balance (AOB), and its structure is shown in Fig. 1 [23].

AFLS’s rules are derived and extracted from given training data. In other words, its parameters can be trained like a neural network approach, but with its structure following the one of a fuzzy logic system. The first layer is the fuzzification layer, and its nodes represent the fuzzy sets used in the antecedent parts of the fuzzy rules. A fuzzification node receives an input and determines the degree to which this input belongs to in the node’s fuzzy set. The outputs of this layer are the values of “Gaussian-shape” membership functions (MF) for the input values.

$$\mu_{F_i^m}(x_i) = \exp \left[-\frac{(x_i - c_i^m)^2}{2(b_i^m)^2} \right] \quad (1)$$

where c_i^m and b_i^m are the centre and spread parameters of the membership function for the i^{th} input and the m^{th} rule. The next layer is the firing strength calculation layer. Since each fuzzy rule’s antecedent part has AND connection operator, the firing strengths are calculated using the product T-norm operator. The most commonly used fuzzy AND operations are intersection and algebraic product. In this case, the multiplication has been used, and the output of this layer has the following form:

$$\mu_m(\bar{x}) = \prod_{i=1}^n \mu_{F_i^m}(x_i), \quad m=1,2,\dots,M \quad (2)$$

where $\mu_{F_i^m}(x_i)$ is the membership value of the i^{th} input of rule m . The number of rules, at this layer, is equal to the number of MFs allocated to each input variable, minimizing thus the problem of excessive number of rules as in the case of TSK-based models. In the classic CA defuzzification scheme, the output of AFLS is a linear combination of the weighted outputs of the normalised rules, as shown in the following equation

$$O_p = \sum_{m=1}^M y_p^m \phi_m(\bar{x}) \quad (3)$$

where y_p^m is considered as the centre of the membership function of the p^{th} output of the m^{th} rule. This layer involves also the normalization rules layer. Each node in this layer calculates the normalized activation strength of each rule by:

$$\varphi_m(\bar{x}) = \frac{\mu_m(\bar{x})}{\sum_{j=1}^M \mu_j(\bar{x})} \quad (4)$$

An alternative defuzzification approach, the area of balance (AOB), has been proposed as an attempt to “match” the accuracy of COA while preserving the simplicity of CA defuzzification method [23]. The widely used COA method generates the centre of gravity of the possibility distribution of a control action. Thus, in the case of a discrete universe, the COA method yields

$$y_{COA} = \frac{\sum_{q=1}^Q \mu_y(y_b^q) y_b^q}{\sum_{q=1}^Q \mu_y(y_b^q)} \quad (5)$$

where Q is a number of quantisation levels of the output, y_b^q is the amount of control output, at the quantization level q and $\mu_y(y_b^q)$ represents its membership value in the output fuzzy set. In general, the higher Q is, the finer y_{COA} will be [21], however the computational cost increases analogous to Q . Since the method provides good performance, its main characteristics, such as centre of gravity and use of the shape of membership function, have been adopted in the design of the proposed defuzzification approach. AFLS’s overall output utilises Kosko’s method with product inference [24].

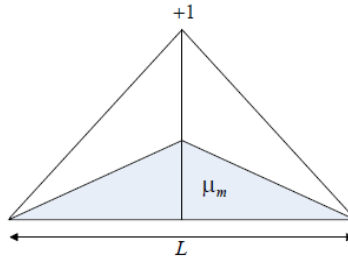


Fig. 2: Triangular shape membership function

The proposed method can be explained by the following mass-less beam example. Let us consider the density (D), which is defined as mass (M) per unit volume (V).

$$D = \frac{M}{V} \quad (6)$$

Under the assumption that we use the same material and all shapes have the same thickness, T, then

$$M = ATD \quad (7)$$

where A is the area. Let us assume for simplicity, that the shape of the membership function used in the consequent part has a symmetric triangular form. The centre of gravity is passed through the halfway point of the base of that shape. For example, if a triangular shape and product-inference as a t-norm are used, then the shape of the consequent part of rule m can be shown as in Fig. 2.

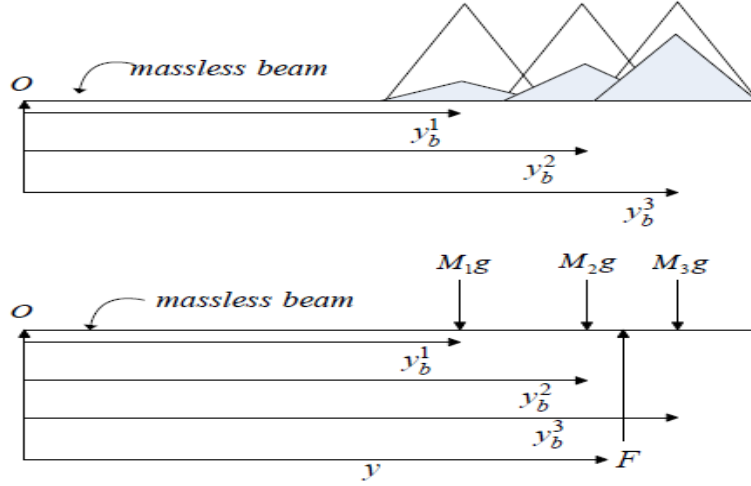


Fig. 3: Consequent fuzzy set placed on mass-less beam

If we could consider the consequent part of each rule placed on the massless beam having the pivot point at origin, then such visualisation is shown in Fig. 3.

Then,

$$F = M_1g + M_2g + M_3g = (M_1 + M_2 + M_3)g \quad (8)$$

For balance

$$Fy = M_1gy_b^1 + M_2gy_b^2 + M_3gy_b^3 = (M_1y_b^1 + M_2y_b^2 + M_3y_b^3)g \quad (9)$$

$$y = \frac{(M_1y_b^1 + M_2y_b^2 + M_3y_b^3)g}{F} = \frac{(M_1y_b^1 + M_2y_b^2 + M_3y_b^3)g}{(M_1 + M_2 + M_3)g} \quad (10)$$

Assume that D in Eq. 7 is the same, thus,

$$y = \frac{(A_1y_b^1 + A_2y_b^2 + A_3y_b^3)TD}{(A_1 + A_2 + A_3)TD} = \frac{(A_1y_b^1 + A_2y_b^2 + A_3y_b^3)}{(A_1 + A_2 + A_3)} \quad (11)$$

The calculation of area A will be depended upon the type of membership function used. Under the assumption of symmetric shape, this method will have comparable capability with the centroid calculation method to approximate the output from the fuzzy set in the consequent part. By utilising the triangle shape as a membership function and the usage of max-product inference, the shaded area A will be derived as:

$$A_m = \frac{1}{2} \mu_m L_m \quad (12)$$

Deriving from Eq. 11, the output, y will be

$$y = \frac{(\mu_1 L_1 y_b^1 + \mu_2 L_2 y_b^2 + \mu_3 L_3 y_b^3)}{(\mu_1 L_1 + \mu_2 L_2 + \mu_3 L_3)} \quad (13)$$

In general form, the calculation of the output O of the AFLS scheme will be

$$O_p = \frac{\sum_{m=1}^M \mu_m L_p^m y_p^m}{\sum_{m=1}^M \mu_m L_p^m} \quad (14)$$

where O_p is the p^{th} output of the network, μ_m , the membership value of the m^{th} rule, L_p^m is the spread parameter of the membership function in the consequent part of the p^{th} output of the m^{th} rule and y_p^m the centre of the membership function in the consequent part of the p^{th} output of the m^{th} rule [23]. The gradient descent learning algorithm scheme has been used to update its various parameters, including those in Gaussian MFs. The update equations for y_p^m , L_p^m , c_i^m and b_i^m are:

$$y_p^m(n+1) = y_p^m(n) + m_y [y_p^m(n) - y_p^m(n-1)] - \eta_y \frac{\partial J}{\partial y_p^m} \bigg|_n \quad (15)$$

$$L_p^m(n+1) = L_p^m(n) + m_L [L_p^m(n) - L_p^m(n-1)] - \eta_L \frac{\partial J}{\partial L_p^m} \bigg|_n \quad (16)$$

$$c_i^m(n+1) = c_i^m(n) + m_c [c_i^m(n) - c_i^m(n-1)] - \eta_c \frac{\partial J}{\partial c_i^m} \bigg|_n \quad (17)$$

$$b_i^m(n+1) = b_i^m(n) + m_b [b_i^m(n) - b_i^m(n-1)] - \eta_b \frac{\partial J}{\partial b_i^m} \bigg|_n \quad (18)$$

where, J_k the objective function is defined as:

$$J_k = \frac{1}{2} \sum_{p=1}^P (O_p(\bar{x}_k) - d_p(\bar{x}_k))^2 \quad (19)$$

with P the number of outputs, d_p the desired response of the p^{th} output and $O_p(\bar{x}_k)$ defined as in Eq. 14.

3. Behavioural models of power amplifier

3.1 Measurement Setup

The experimental test-bed for measurement purposes is illustrated at Fig. 4. Setup consists of signal generator Agilent MXG N5182A, the power amplifier MIMIX CFH2162-P3 as device under test (DUT), with 1dB compression point at 27 dBm output power level and 13dB nominal gain, one vector signal analyzer (VSA) Agilent E4406A and one PC with suitable software [25]. The required software utilized for these measurements, involve MATLAB, for creating baseband signals for signal generator, the Agilent Signal Studio Toolkit, for assistance in downloading process and finally the Agilent Distortion Suite 89604A for capturing signals from VSA to PC. An LTE 1.4 MHz 64 QAM signal, created in MATLAB environment, was used as an excitation of the DUT. This signal was then downloaded to the MXG via a General Purpose Interface Bus (GPIB), whereas it was up-converted to RF frequency of 880 MHz. The specific “download” procedure was performed via the usage of Agilent Signal Studio Toolkit. Subsequent measurements were carried out at power level of 1dB compression point. Signal at the output of

MXG was passed through DUT. An attenuator of 30dB was added after the amplifier in order to make the signal suitable for VSA. Signal was down-converted by VSA and captured on the PC using the Distortion Suite 89604A. The measured signals were transferred from distortion suite to MATLAB for further processing and analysis.

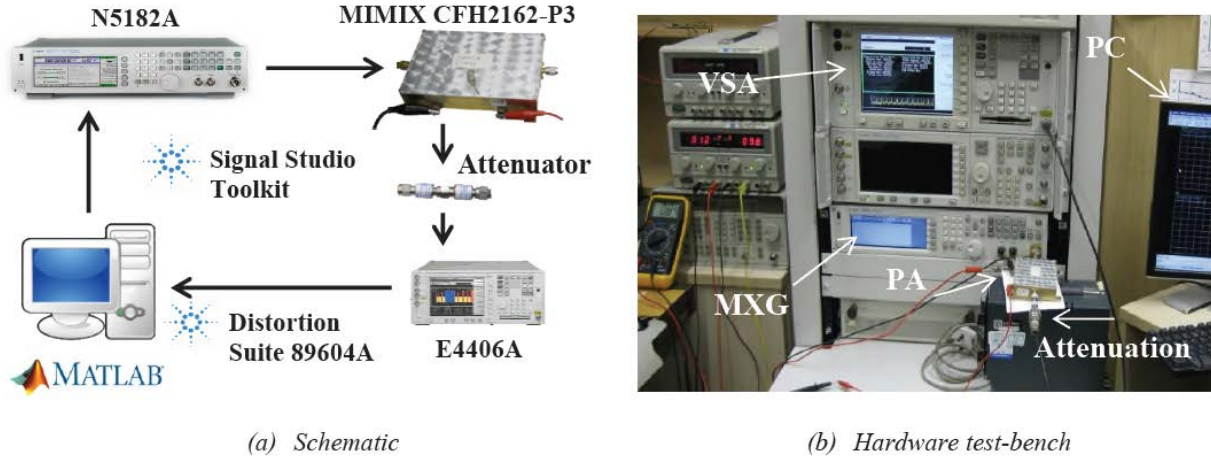


Fig. 4: Experimental setup

3.2 Intelligent-based PA Forward Modelling

In order to achieve maximum efficiency, PA should be driven near the saturation region, but since the orthogonal frequency-division multiplexing (OFDM) signals have high peak-to-average power ratio (PAPR), PA will cross over to the nonlinear region causing serious in-band distortions as well as adjacent channel interference (ACI) with spectrum re-growth (out-of-band distortion) in the transmitted signal [12], [26]. Nonlinear effects can be minimised by operating PA preferably in the linear region, but in this case, efficiency has to be compromised, thus the performance of an OFDM system will have to be degraded [27].

DPD techniques are widely used as the most cost-effective solutions for compensation of distortion effects [28]. Using DPD techniques, an inverse model of PA's behaviour could be derived, and in cascade with PA could make the system with more or less improved linearization, depending on method. PA behavioural modelling usually requires a model that can extract amplitude and phase information from modulated complex waveforms.

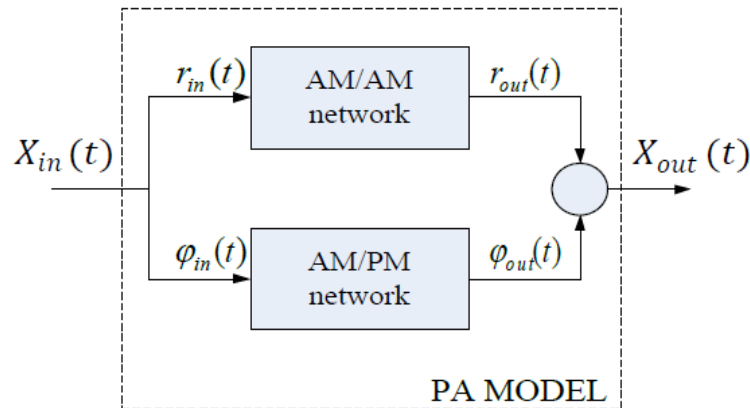


Fig. 5: Conventional NN topology for PA modelling

The most basic proposed learning-based structure is a single-input single-output feed-forward NN utilizing complex input/output [29]. Such approach is illustrated at Fig. 5. One NN has been utilised to model the dynamic AM/AM characterization, while the other one to represent the AM/PM characteristic of DUT. The training dataset for AM/AM NN includes the amplitude of the PA input signal as the network training input, and the amplitude of the PA output signal as the desired output. The other network, AM/PM, is associated with the training dataset composed of the amplitude of the PA input signal as the network training input, and the phase difference between the output and input signals of the PA as the desired output. Once both networks are trained, they are put in a parallel, creating a model of DUT.

A more effective approach, called the time-delay neural network (TDNN) is shown in Fig. 6, which takes advantage of the easy availability of the in-phase (I) component and quadrature (Q) component of the modulated waveform in the baseband, thereby saving pre and post-processing activities and can be used as a common feed-forward NN with two inputs and two outputs [30]. In order to incorporate the memory effects existing in PA behaviour, the TDNN scheme incorporates related time-delayed inputs.

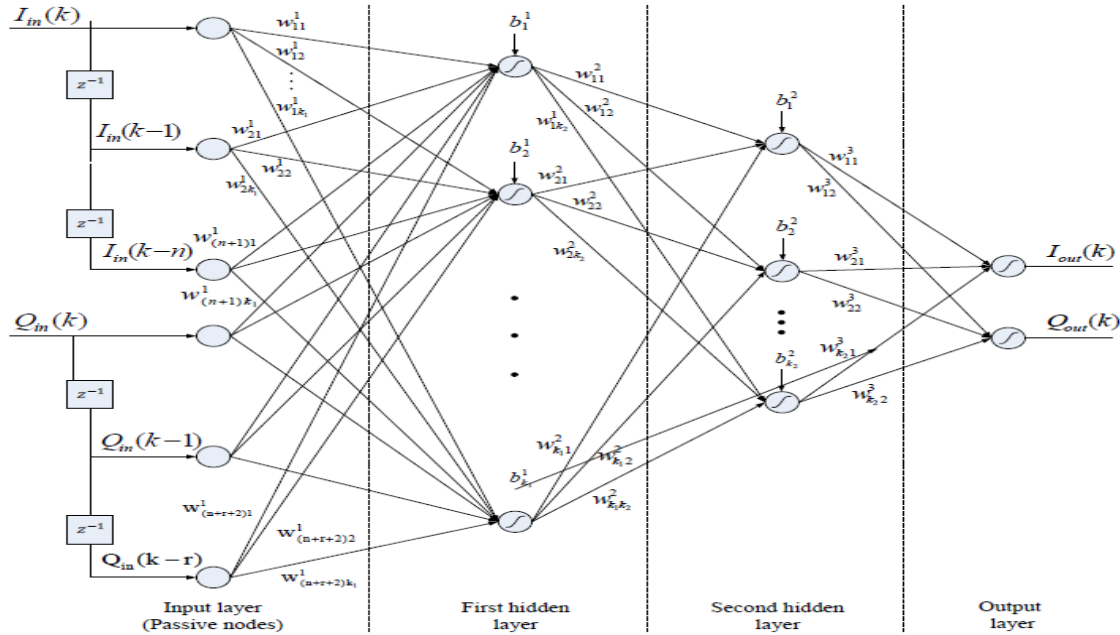


Fig. 6: TDNN scheme modelling

The TDNN model contains $n + r + 2$ input variables, which are related with the current in-phase component and its n time-delayed samples as well as the current quadrature component with r of its time-delayed samples. In general, although the number of n and r delay taps can be assigned with any arbitrary integer values, for this specific type of problem, these numbers have been chosen to be equal, as in-phase and quadrature components are affected equally by PA's nonlinearity. The problem of order determination for nonlinear dynamic systems is still not satisfactorily solved and very little research seems to be devoted in this research area [20]. It is common practice to select the dynamic order of the model by a combination of trial and error and prior knowledge about the process. Obviously, if oscillatory behaviour is observed, the process must be at least of second order ($n, r \geq 1$). TDNN's

internal structure consists of two hidden layers and the classic sigmoid activation function has been adopted for both hidden and output layers. The complex input signal in the time sample k has the Cartesian form

$$X_{in}(k) = I_{in}(k) + jQ_{in}(k) \quad (20)$$

where $I_{in}(k)$ and $Q_{in}(k)$ are real values. The network's outputs are then defined as:

$$\begin{bmatrix} I_{out}(k) \\ Q_{out}(k) \end{bmatrix} = \begin{bmatrix} f_I(I_{in}(k), I_{in}(k-1), \dots, I_{in}(k-n), Q_{in}(k), Q_{in}(k-1), \dots, Q_{in}(k-r)) \\ f_Q(I_{in}(k), I_{in}(k-1), \dots, I_{in}(k-n), Q_{in}(k), Q_{in}(k-1), \dots, Q_{in}(k-r)) \end{bmatrix} \quad (21)$$

In this paper, the same memory-based modelling approach provided by Eq. 21, has been adopted also in the proposed AFLS-based model. Both input and output signals have been normalized down to range [0, 1]. Final AFLS's input vector consisted of present and past inputs of I and Q components of the complex input signal (i.e. $\{I(k), Q(k), I(k-1), Q(k-1)\}$). Optimal structure in terms of number of delay taps and number of memberships / rules has been achieved through a trial and error procedure. The number of membership functions per input in the fuzzification layer, provides a localised learning in NF systems, and generally that leads to smaller input vector requirements, for the case of dynamic systems [31]. Fig. 7 illustrates the AFLS model tailored to this case study.

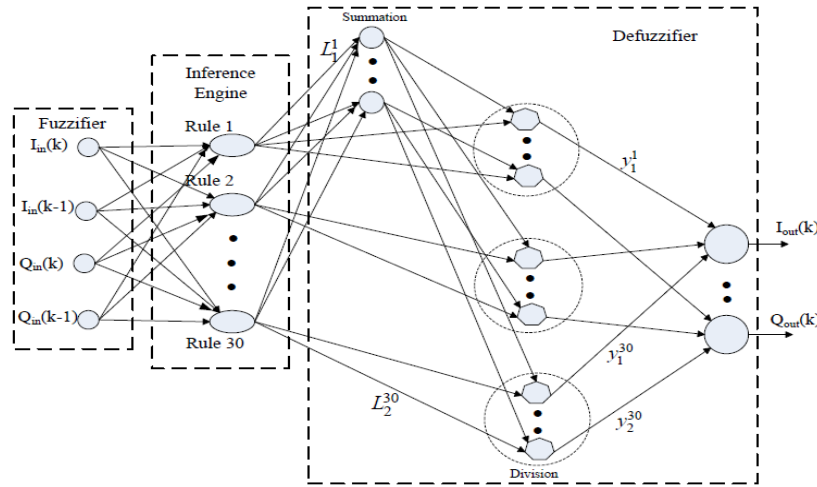


Fig.7: AFLS for PA's forward modelling

During that procedure, the performance of AFLS structure was observed through the normalized mean square error (NMSE) that combines errors of complex signal components, I and Q :

$$NMSE_{dB} = 10 \log_{10} \left(\frac{\sum_{n=1}^K (I_{meas}(n) - I_{est}(n))^2 + (Q_{meas}(n) - Q_{est}(n))^2}{\sum_{n=1}^K (I_{meas}(n)^2 + Q_{meas}(n)^2)} \right) \quad (22)$$

where I_{meas} and Q_{meas} were the measurements of baseband output signal, while I_{est} and Q_{est} the normalised to range [0-1] AFLS's output signals. NMSE usually is considered as a regular tool in such optimisation processes in wireless technology [17]. After many trials, it has been found that only 30 rules were necessary for the proposed AFLS model to achieve an acceptable performance for this particular experiment. The number of membership functions for each input variable was directly associated to the number of rules, hence, each input signal was “distributed” through Gaussian functions with different centres and widths to every rule node via a product operator.

The learning rates for training of each parameter were set to 0.008, while momentum parameters to 0.55. The initial centre parameters of Gaussian membership functions at the fuzzification layer were equally distributed in range [0, 1], while the initial spread parameters of these membership functions to 0.2. Finally, the initial values of centres of the membership functions y_p^m in the defuzzification part were equally distributed in the range [0, 1] while the initial values of spread parameters L_p^m to 0.65. The training dataset consisted of the first 30,000 (30K) patterns of the specific signal, while the remaining 100K patterns were used as a testing dataset.

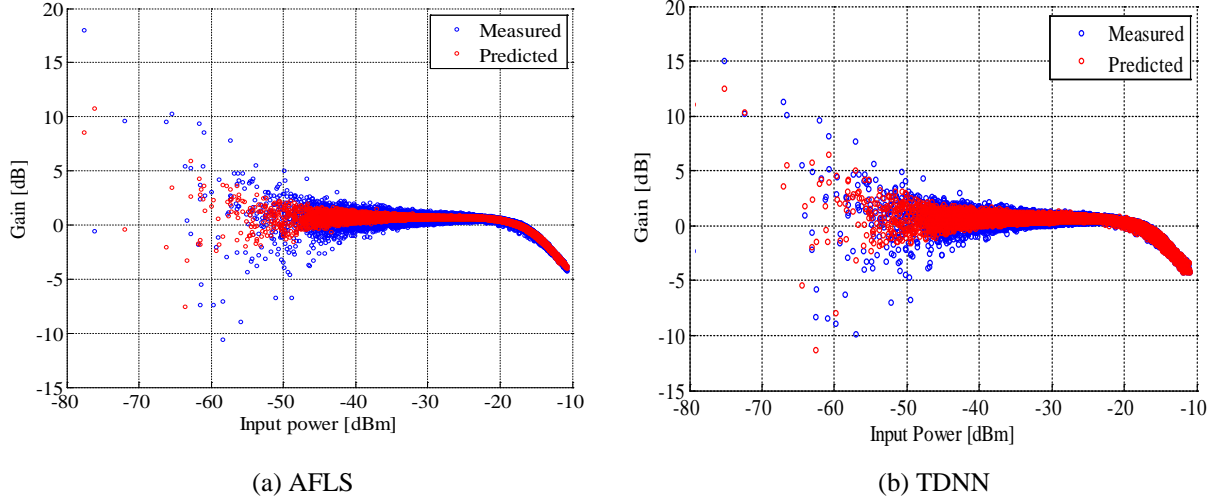


Fig. 8: PA gain for the LTE 1.4 MHz 64 QAM signal (Forward Models)

Evaluation of the behavioural model performance of PAs exhibiting nonlinear memory effects is a very critical task. Accordingly, additional metrics need to be considered to evaluate the accuracy of a behavioural model in the predicting the in-band and out-of-band distortion of DUT behaviour. In cases where the out-of-band performance of the PA is of more importance, the adjacent channel power ratio (ACPR) can be used, which is defined as the ratio between the signal power of the adjacent channel and the power of the measured output signal within the channel. The ACPR can be calculated for both the lower and upper adjacent channels. In a logarithmic representation ACPR can be presented as

$$ACPR[dB] = P_{adj} [dBm] - P_{fund} [dBm] \quad (23)$$

where P_{fund} denotes overall power integrated in the main channel, expressed in dBm , while P_{adj} is defined as

$$P_{adj} [dBm] = 10 \log_{10} (P_{adjL} + P_{adjU}) \quad (24)$$

where P_{adjL} and P_{adjU} are powers integrated in adjacent channels. Peak to average power ratio (PAPR) is defined as the ratio between the maximum instantaneous power in time and the average power of the output signal

$$PAPR[dB] = 10 \log_{10} \frac{\max_{0 \leq n \leq N-1} |x(n)|^2}{P_{avg}} \quad (25)$$

Single-carrier systems are not exposed to the impairments in the front-end, as the OFDM transceivers. The high PAPR problem associated with multicarrier signals is one of the principal impairments in the implementation of

OFDM systems. The linear operation of the PA over a large dynamic range increases the implementation cost and reduces the power efficiency. Output signals were normalised down to the main channel level of input signal, in order to display the distortion, and subsequently to illustrate the level of distortion suppression.

The error vector magnitude (EVM) is a common figure of merit for assessing the quality of digitally modulated signals. It is defined as the difference between magnitudes of the ideal reference signal and the measured output (transmitted) signal, after the compensation in time, amplitude, frequency, phase and DC offset.

$$EVM [\%] = \frac{\sum_{i=1}^n \sqrt{[I(i) - \hat{I}(i)]^2 + [Q(i) - \hat{Q}(i)]^2}}{\sum_{i=1}^n \sqrt{[I(i) + Q(i)]^2}} * 100 \quad (26)$$

where $I(i)$, $\hat{I}(i)$ and $Q(i)$, $\hat{Q}(i)$ are the, normalized to range $[0, 1]$, reference and measured I , Q components respectively. This particular metric is widely utilised by microwave engineers as it contains information about amplitude and phase errors in the signal [28]. NMSE for testing the “training” and “testing” datasets were -43.54dB, and -42.08dB respectively for the specific LTE 1.4 MHz 64 QAM signal. PAPR for the input signal has been calculated to 10.21dB. Figs. 8a and 9 illustrate AFLS’s modelling performance of the PA, using the LTE 1.4 MHz 64 QAM signal at 880MHz as centre frequency. Fig. 8a represents the PA gain at 1dB compression point.

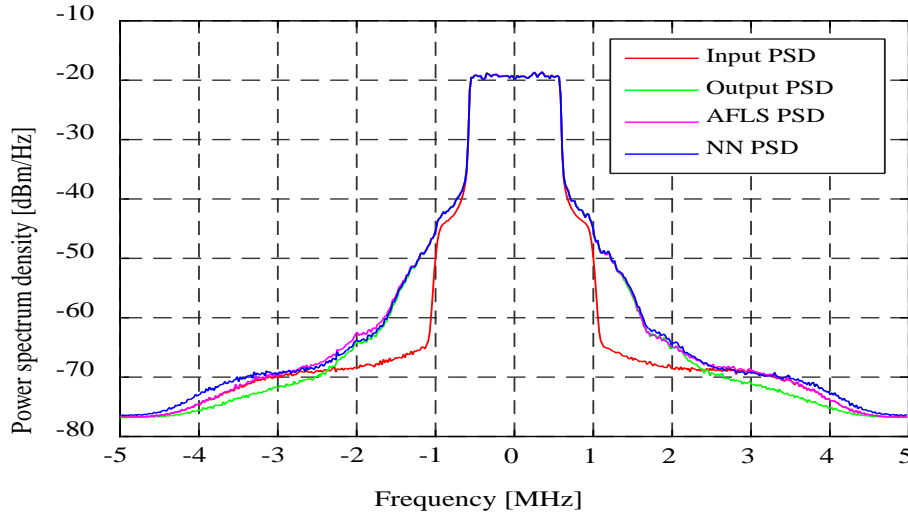


Fig. 9: Power spectrum density of the AFLS/TDNN model outputs vs. the real PA (Forward Models).

As it is shown from Fig. 9, PA’s AFLS-based forward model seems to be very accurate. Table 1 provides the performance indicators for the measured input and output signals, as well as AFLS’s output signal. Fig. 8a reveals very good matching results for the PA nonlinear gain, while almost identical spectra of the real PA output and AFLS’s output confirm model’s accuracy. Such performance can be also verified through EVM and ACPR performance parameters, while an excellent NMSE provides additional evidence of such good modelling structure.

Although AFLS utilises the standard gradient descent method for learning purposes, training process was performed very fast (i.e. approx. 90 min), using a PC with an i7-2630QM CPU. The main reason for achieving such fast training process is due to the “structural design” of neurofuzzy systems. All NF systems follow the so-called

localised learning compared to MLP's global learning concept [32]. Such localised learning is achieved through the use of Gaussian based functions in the fuzzification part.

Forward Model	ACPR [dB] ±1.4 MHz offset		EVM [%]
	LOW	HIGH	
PA input	-36.19	-34.28	/
Real PA	-30.89	-30.00	8.66
AFLS model (training dataset)	-30.84	-29.98	8.54
AFLS model (testing dataset)	-30.82	-30.11	8.53
TDNN model (testing dataset)	-30.85	-30.11	8.58

Table 1: Forward model performance

In parallel with the proposed AFLS architecture, a TDNN structure has been attempted, using time-delayed inputs for I and Q components. After a few trials, a suitable TDNN was constructed with an input vector consisted of eight variables, $\{I(k), Q(k), I(k-1), Q(k-1), I(k-2), Q(k-2), I(k-3), Q(k-3)\}$ with two hidden layers (with 20 and 12 nodes respectively). The gradient descent learning method has been also utilized to this case, while learning rate and momentum parameters have been set to 0.1 and 0.2 respectively. The NMSE for the testing dataset was calculated to -38.28 dB, while the related PAPR index to 9.92dB for the same LTE 1.4 MHz 64 QAM signal. Following Table 1, TDNN had a similar to AFLS performance in terms of ACPR and EVM, however, such result has been achieved with a high computational cost (i.e. 24 hours training time and enormous network structure). In theory, TDNN could be simplified by adopting the one-hidden layer approach, however practice has revealed that in real applications, always a two-hidden layer strategy outperforms in terms of accuracy the single one-hidden-layer approach. Figs. 8b and 9 illustrate TDNN's modelling performance of the PA, using the LTE 1.4 MHz 64 QAM signal at 880MHz as centre frequency utilising the same training and testing dataset as in the case of AFLS.

In terms of complexity, the adopted TDNN structure required 418 weights to be optimised via training process, while AFLS model only 360 parameters (including centres/spreads and weights at fuzzification and defuzzification parts respectively). However, complexity could be also associated with the amount of time needed for training. As TDNN training is based on a global learning approach compared to the local learning for the case of AFLS, more than 50,000 epochs were required for TDNN to achieve an acceptable performance. In the case of AFLS, however only 4,000 epochs required, even using the same gradient descent learning algorithm.

The choice on not utilising ANFIS model in this case study is further justified by the exponential number of rules and parameters required for designing such model. With the same number of inputs, as in the AFLS case, and with three minimum fuzzy memberships per input, 81 fuzzy rules are needed for an ANFIS architecture, which is considered as an operation with high computational cost. As the defuzzification section follows the classic TSK scheme, 405 and 24 parameters are required for the consequent and premise section of this architecture.

3.3 Intelligent-based PA Inverse Modelling

The principle of pre-distortion is to distort the PA input signal by an additional device called a pre-distorter (PD), whose characteristics are the inverse of those of the amplifier. In our case, the main aim is the development of a learning-based model that inverts PA's nonlinearities/dynamics. Such an inverse model could be placed in cascade with the PA, in order to provide an overall linearization of the transmission process. An AFLS-based inverse model structure has been developed using the same principles as the forward model, however the number of rules was increased to 40, while the input vector required additional time-delayed variables, increasing to six inputs, $\{I(k), Q(k), I(k-1), Q(k-1), I(k-2), Q(k-2)\}$. Learning and momentum rates were kept the same as for the forward model case. The training procedure has been performed, with the reverse input-output dataset. The input signal was considered as the desired output, while the output baseband signal was the training dataset. Similarly to the previous case, 30K and 100K patterns were used for training and testing respectively.

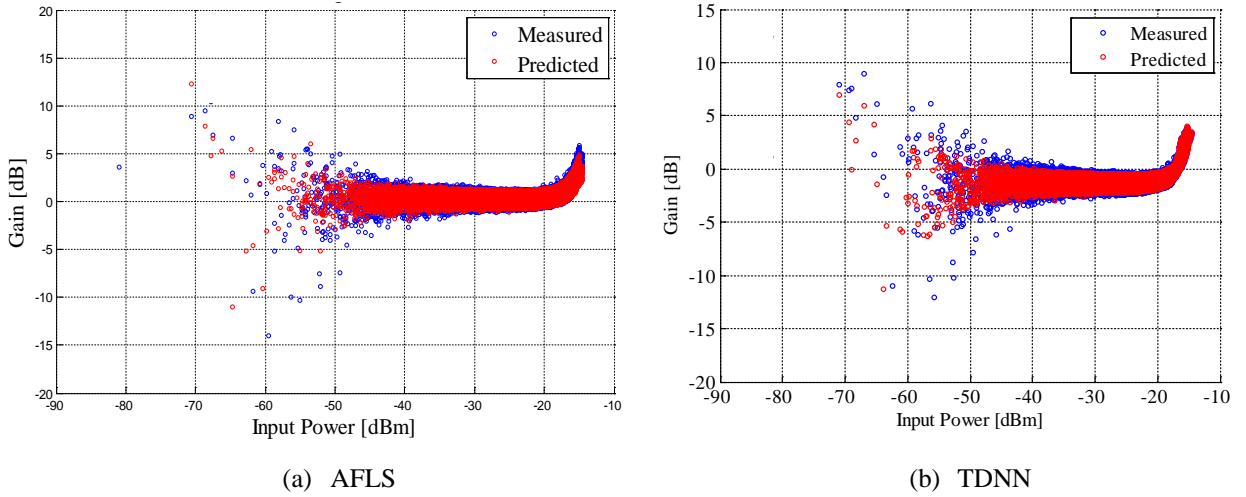


Fig. 10: PA gain for the LTE 1.4 MHz 64 QAM signal (Inverse Models)

With the increased number of input variables as well as rules, 6000 epochs were required to adequately train the specific AFLS inverse model. NMSE for testing the “training” and “testing” datasets were -44.27dB, and -44.15dB respectively for the specific LTE 1.4 MHz 64 QAM signal. Figs. 10a and 11 show AFLS’s “inverse” modelling performance of the PA, using the LTE 1.4 MHz 64 QAM signal at 880MHz as centre frequency. It has to be noted, that labels “Input Power” and “Input PSD” at Figs. 10 and 11 respectively, are actually the PA’s output, as according to identification theory, in inverse modelling, PA’s output represents the input in the inverse model. In addition, Table 2 presents the statistical performances for the inverse modelling development. A more detailed inspection at the spectra of the models’ output signals revealed some mismatches within frequency bands. The scattered samples on lower power levels were a consequence of the out-of-band model behaviour, which did not track the DUT perfectly. Spectra of the models’ output signals had also mismatches in the beginning and ending frequency regions. The explanation for that was related to the frequency range of the measurement. It would be preferable that the measurement of the signals be in range five times bigger than the signal occupied bandwidth.

Additionally, mismatches detected in a conversion through the VSA and distortion suite. Filtering brings an additional transfer part in the system, thus changing the overall transfer characteristics, different from the one in the main channel [25]. However, very good modelling results for the testing dataset have been produced, and ACPR and EVM metrics simply verify the outcome.

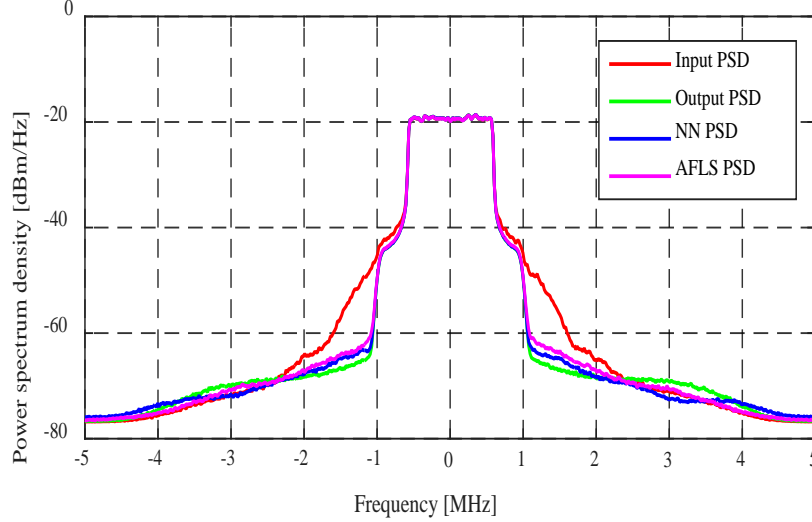


Fig. 11: Power spectrum density of the AFLS/TDNN model outputs vs. the real PA (Inverse Models).

Similarly to the AFLS case, a TDNN has been developed for the “inverse” modelling problem. Two hidden layers with 15 and 6 nodes have been adopted for this case, while the input vector increased to 10 inputs in order to provide an acceptable performance (*i.e.* $\{I(k), Q(k), I(k-1), Q(k-1), I(k-2), Q(k-2), I(k-3), Q(k-3), I(k-4), Q(k-4)\}$). The NMSE for testing dataset was calculated as -36.67dB for the LTE 1.4 MHz 64 QAM signal. Good performance as indicated from Table 2 was achieved however through an extensive training time of 50,000 epochs, lasting more than 40 hours. Supervised learning of a NN is viewed as surface-fitting process in a high-dimensional surface. The dimensions of such surface are direct related to the number of NN’s input variables [20]. In the case of inverse modelling, TDNN required 10 input variables, compared to 6 inputs for the case of AFLS. Hence, the difficulty and time required to obtain an accurate approximation surface for the case of TDNN has been significantly increased. Fig. 11 illustrates also TDNN’s “inverse” modelling performance of the PA, using the LTE 1.4 MHz 64 QAM signal at 880MHz as centre frequency.

Inverse Model	ACPR [dB] ± 1.4 MHz offset		EVM [%]
	LOW	HIGH	
AFLS model (training dataset)	-35.56	-33.51	2.38
AFLS model (testing dataset)	-35.63	-33.76	2.42
TDNN model (testing dataset)	-36.01	-34.15	1.47

Table 2: Inverse model performances

In TDNN models, all normalized inputs are fed to the hidden layer, while in the case of AFLS each input is fuzzified / decomposed through Gaussians membership functions. Both inverse models share the same learning training algorithm, i.e. the gradient descent method. In terms of complexity, the adopted inverse TDNN structure required 275 weights to be optimised via training process, while AFLS model 640 parameters (including centres/spreads and weights at fuzzification and defuzzification parts respectively). As the number of these membership functions is equal to the numbers of rules, this architecture has however advantages over the classic ANFIS neuro-fuzzy model. In fact, an inverse ANFIS model utilising the same number of input variables as in the case of AFLS, would require 729 fuzzy rules (under the assumption that minimum three membership functions were used per input). Such enormous structure is practically not efficient for any real application. The increased number of Gaussian membership functions in AFLS-based inverse model, although increases the number of parameters to be optimised, maintains however the required number of rules at low level.

4. Power Amplifier Pre-distortion using Inverse Model

In the pre-distortion concept, a pre-distorter block is placed before the amplifier block, and this pre-distorter should exhibit a behaviour which is the inverse of the amplifier's nonlinear behaviour, so that the two blocks together ultimately behave linearly. Fig. 12 shows a layout of a pre-distortion system. The pre-distorter should in fact have a transformation function to compensate for the amplifier's nonlinear behaviour. An ideal pre-distortion response is the inverse response function of the RF transmitter [17].

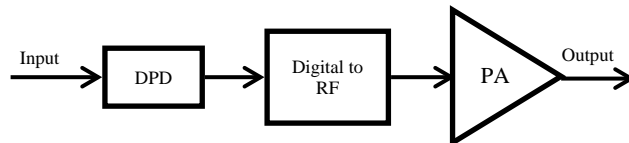


Fig. 12: Generic Layout of a Pre-distortion System

Initially, both inverse models (AFLS and TDNN) were evaluated as potential pre-distorters in cascade with the “forward” learning-based models of the DUT, in a simulation experiment. Then, pre-distorter systems were tested for the LTE 1.4 MHz 64 QAM signal on the real DUT (PA MIMIX CFH2162-P3). The scenario was to pass first an input signal to the pre-distorter unit, (PA's inverse model) and then through the PA to obtain a linearly gained signal at the output.

The spectra of the input signal, output signal and signal at the output of the cascade DPD+PA are illustrated at Fig. 13. Fig. 13a illustrates the performance of the DPD on the AFLS forward model which resembles the real PA, while in Fig. 13b, the proposed DPD scheme is applied on the real PA. In these two experiments, the input signal (LTE 1.4 MHz 64 QAM) was used as the input of the pre-distorter, which is the first block in the DPD system. AFLS inverse model has been utilized in this case as the DPD component. For the simulation pre-distortion case, spectral re-growth improvement for the LTE 1.4 MHz 64 QAM signal was 19.58dB for the upper band, and 18.50dB for the lower band. Spectral re-growth improvement was calculated as difference between power levels of distorted output signal and signal at the output of system with DPD, at ± 1.2 MHz offset. The results acquired from testing the pre-distorter in real environment, without using the AFLS model for the device under test, revealed again

an improvement of the spectral re-growth at the level of 15.18dB for the upper band, and 15.00dB for the lower band.

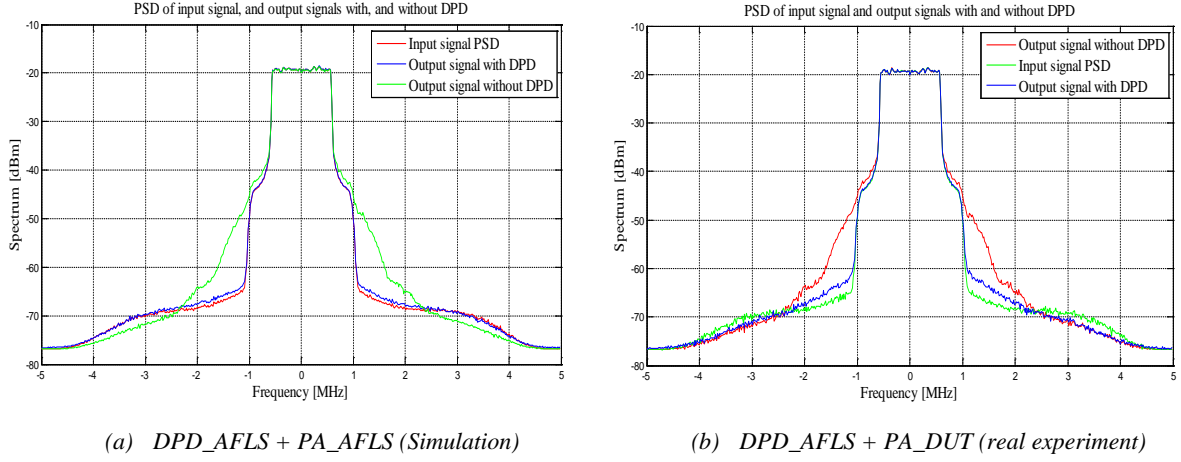


Fig. 13: DPD Power spectrum densities at 880 MHz for the LTE 1.4 MHz 64 QAM signal (AFLS case)

These spectra have shown that it is possible to decrease the distortion coming from the PA using the proposed intelligent based scheme. This is also observable through the performance parameters, ACPR and EVM. As can be seen from Table 3, ACPRs were very close to ACPR of the input signal and EVM was very small. That means that the output of system with a pre-distorter was almost completely linear transformation from the original input.

AFLS - case	ACPR [dBc] (Simulation) ±1.4 MHz offset		ACPR [dBc] (Real PA) ±1.4 MHz offset		EVM [%]
PA input	LOW -36.66	HIGH -34.76	LOW -36.24	HIGH -34.32	/
PA output without DPD	-26.98	-26.61	-30.91	-30.02	8.68
PA output with DPD	-36.49	-34.64	-36.06	-34.18	1.32

Table 3: Pre-distortion System Performance using AFLS model

PA's gain in the DPD+PA configuration differs from its counterpart without using pre-distortion as shown from Fig. 14, while Fig. 15 reveals an almost perfect matching with a linear system regarding AM/AM and AM/PM dynamical characteristics.

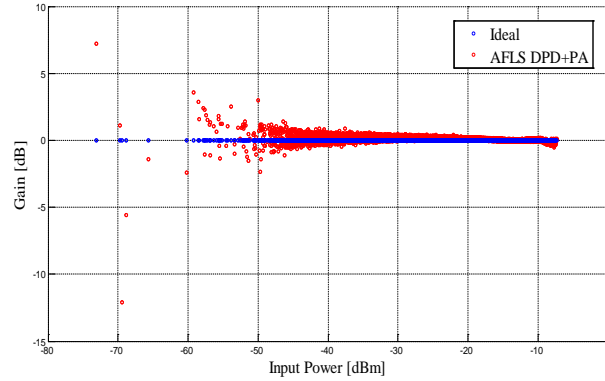


Fig. 14: PA's Gain in a DPD+PA configuration

Unlike existing predistortion techniques which utilise two separate ANFIS networks for the power amplifier linearization, the presented method utilised a compact structure simpler than earlier NF approaches.

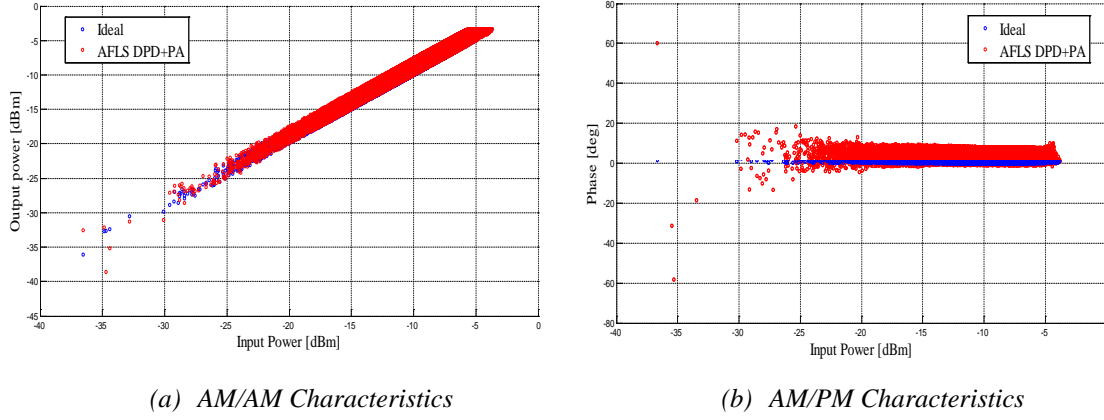


Fig. 15: DPD+PA characteristics

Similarly to the case of AFLS, the implemented TDNN inverse model has been also utilized as part of DPD configuration. The spectra of the input signal, output signal and signal at the output of the cascade DPD+PA using TDNN are illustrated at Fig. 16. Fig. 16a illustrates the performance of the DPD on the TDNN forward model which resembles the real PA, while in Fig. 16b, the DPD-TDNN scheme is applied on the real PA.

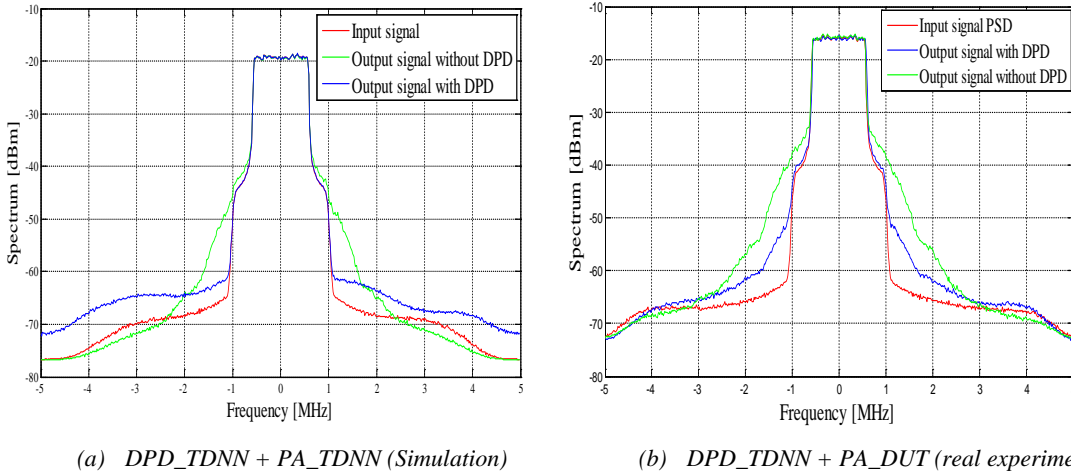


Fig. 16: DPD Power spectrum densities at 880 MHz for the LTE 1.4 MHz 64 QAM signal (TDNN case)

For the simulation pre-distortion case, spectral re-growth improvement for the LTE 1.4 MHz 64 QAM signal was 11.76dB for the upper band, and 11.44dB for the lower band. The results acquired from testing the pre-distorter in real environment, without using the TDNN model for the device under test, revealed again an improvement of the spectral re-growth at the level of 10.27dB for the upper band, and 11.14dB for the lower band. As can be also seen from Table 4, ACPRs were close to ACPR of the input signal and EVM was very small. That means that the output of system with a pre-distorter was an acceptable linear transformation from the original input.

The comparison of the presented techniques was for the DUT at frequency of 880MHz for the same LTE signal with 1.4 MHz bandwidth. Adjacent channels for measurement of the ACPR performance were at 1.4 MHz offset with

bandwidth of 1.08 MHz. The trainings of the pre-distorters were carried out on the same PC, in order to have the same base for comparison. This 1.08 MHz channel bandwidth was the real occupied bandwidth of the used LTE 1.4 MHz 64 QAM signal.

TDNN - case	ACPR [dBc] (Simulation) ±1.4 MHz offset		ACPR [dBc] (Real PA) ±1.4 MHz offset		EVM [%]
	LOW	HIGH	LOW	HIGH	
PA input	-36.22	-34.30	-36.69	-34.76	/
PA output without DPD	-30.90	-30.01	-26.98	-26.61	8.66
PA output with DPD	-35.73	-33.85	-33.73	-32.50	1.37

Table 4: Pre-distortion System Performance using TDNN model

Table 5 illustrates a comparison of the intelligent based schemes utilised for DPD pre-distortion. EVM results reveal that predistortion process was successful for both methods. The ACPR difference and spectral re-growth show a very good performance of the AFLS method in the proposed inverse-based DPD system. It has to be mentioned that ACPR is in logarithmic scale, thus AFLS and TDNN performances are not considered as equivalent. In addition, distortion suppression (DS) is much improved using the AFLS scheme both in simulation but also in the experimental case. DS is calculated by checking the differences of power levels at ± 1.2 MHz offset of input signal /output without DPD and input signal/output with DPD. Furthermore, the time required for the AFLS training with the same computer was much smaller than for the neural network. As a conclusion, distortion suppression and training time are the main parameters that point out AFLS's superiority over TDNN.

	AFLS (Real PA)	TDNN (Real PA)
ACPR difference between linearly gained signal and output of DPD+PA system [dB] ± 1.4 MHz offset	LOW -0.18	LOW -2.96
	HIGH -0.14	HIGH -2.26
EVM [%]	1.32	1.37
Distortion suppression (DS) [%]	Simulation: 93.3	Simulation: 78.23
	Real: 84.2	Real: 48.4
Structure of pre-distorter	<ul style="list-style-type: none"> • 6 inputs • 40 rules 	<ul style="list-style-type: none"> • 10 inputs • 15 neurons in first hidden layer • 6 neurons in second hidden layer
Training time	~2h	>24h

Table 5: Comparative overview of used DPD techniques for LTE 1.4 MHz 64 QAM signal

5. Conclusions

Linearization of power amplifiers is essential considering requirements of nowadays communication systems. A novel technique for modelling and linearization based on a neuro-fuzzy approach has been described. A detailed

structure explanation was presented as well as the training procedure and the update of system parameters. This technique was explored for the LTE 1.4 MHz 64 QAM signal at 880 MHz as centre frequency driven to the MIMIX CFH2162-P3 power amplifier. Models for forward and inverse characteristics were validated through NMSE and matching spectra. Pre-distortion of the device under test was best shown through the spectral re-growth improvement. A comparative study utilizing a TDNN-based DPD scheme has been carried out, which proved the superiority of AFLS scheme in terms of training time and the improvement in spectral re-growth. Future work will be focused on improving AFLS model by incorporating a pre-processing clustering stage. Such initiative will eventually reduce complexity of AFLS network, thus reducing more the training time and making system even more convenient for use in real-life systems. Furthermore, investigation of applicability of the proposed methods on different physical devices, like Field Programmable Gate Array (FPGA) evaluation platform, will be carried out.

References

1. P. B. Kenington (2000) High-Linearity RF Amplifier Design. Boston, MA: Artech House.
2. F. H. Raab, P. Asbeck, S. Cripps, P. B. Kenington, Z. B. Popovic, N. Potheary, J. F. Sevic, and N. O. Sokal (2002) Power Amplifiers and Transmitters for RF and Microwave. IEEE Transactions on Microwave Theory and Techniques 50(3): 814-826.
3. R. N. Braithwaite and A. Khanifar (2013) High efficiency feedforward power amplifier using a nonlinear error amplifier and offset alignment control. IEEE MTT-S Int. Microwave Symposium, Seattle, WA, 1-4
4. Y. Kim, Y. Yang, S. H. Kang, and B. Kim, (1998) Linearization of 1.85 GHz amplifier using feedback predistortion loop. IEEE MTT-S Int. Microw. Symp. Dig., Baltimore, MD, 1675-1678
5. J. Yi, Y. Yang, M. Park, W. Kang, and B. Kim (2000) Analog Predistortion Linearizer for High-Power RF Amplifiers. IEEE Trans. Microwave Theory Tech. 48(12): 2709-2713.
6. E. G. Jeckeln, F. M. Ghannouchi, and M. A. Sawan (2004) A new adaptive predistortion technique using software-defined radio and DSP technologies suitable for base station 3G power amplifiers. IEEE Trans. Microw. Theory Tech. 52(9) : 2139-2147.
7. A. Zhu, M. Wren, and T. J. Brazil, (2003) An efficient volterra-based behavioral model for wideband RF power amplifiers. IEEE MTT-S Int. Microw. Symp. Dig. 2 : 787-790
8. Jardin P, Baudoin G. (2007) Filter lookup table method for power amplifier linearization. IEEE Trans Veh Technol. 56: 1076-1087
9. H. C. Ku, M. D. McKinley, and J. S. Kenney (2002) Quantifying memory effects in RF power amplifiers. IEEE Transactions on Microwave Theory and Techniques 50(12) : 2843-2849
10. R. Morgan, M. Zhengxiang, L. Kim, M. G. Zierdt, and I. Pastalan (2006) A generalized memory polynomial model for digital predistortion of RF power amplifiers. IEEE Trans. Signal Processing, 54(10) : 3852-3860.
11. S. Boumaiza, and F. Mkadem, (2009) Wideband RF Power Amplifier Predistortion using Real-Valued Time-Delay Neural Networks. European Microwave Conference, EuMC2009, 1449-1452
12. V.P. Gil Jiménez, Y. Jabrane, A. García Armada, B. Ait Es Said, and A. Ait Ouahman (2011) High Power Amplifier Pre-Distorter Based on Neural-Fuzzy Systems for OFDM Signals. IEEE Transactions on Broadcasting, 57(1) : 149-158
13. N. Naskas and Y. Papananos (2002) Adaptive baseband predistorter for radio frequency power amplifiers based on a multilayer perceptron. 9th Int. Electron.. Circuits, Syst. Conf. 3(11) :1107-1110

14. K. C. Lee, and P. Gardner, (2006) Adaptive neuro-fuzzy inference system (ANFIS) digital predistorter for RF power amplifier linearization, *IEEE Trans. Veh. Technol.*, 55(1) : 43–51
15. J. Zhai, J. Zhou, L. Zhang, J. Zhao, and W. Hong, (2009) The Dynamic Behavioral Model of RF Power Amplifiers with the Modified ANFIS. *IEEE Transactions on Microwave Theory and Techniques*, 57(1) : 27-35
16. Jang JSR, Sun CT, Mizutani E. (1997) *Neuro-fuzzy and soft computing: a computational approach to learning and machine intelligence*, Prentice-Hall
17. M. Rawat, K. Rawat, F.M. Ghannouchi, (2010) Adaptive Digital Predistortion of Wireless Power Amplifiers/Transmitters Using Dynamic Real-Valued Focused Time-Delay Line Neural Networks. *IEEE Transactions on Microwave Theory And Techniques*, 58(1) : 95 -104
18. T. Liu, S. Boumaiza, and F. Ghannouchi, (2004) Dynamic behavioural modeling of 3G power amplifiers using real-valued time-delay neural networks. *IEEE Trans. Microw. Theory Tech.*, 52(3) : 1025–1033
19. L.H. Tsoukalas, R.E. Uhrig, (1997) *Fuzzy and Neural Approaches in Engineering*, Wiley
20. O. Nelles, (2000) *Nonlinear system Identification: From Classical Approaches to Neural Networks and Fuzzy Models*. Springer, Berlin
21. J. Mendel (2001) *Uncertain rule-based fuzzy logic systems*, Prentice Hall
22. L. X. Wang (1994) *Adaptive Fuzzy Systems and Control*, Prentice Hall, Inc., Englewood Cliffs
23. V.S. Kodogiannis, I. Petrounias and J.N. Lygouras (2008) Intelligent Classification using Adaptive Fuzzy Logic Systems. *IS2008 – IEEE International Conference on Intelligent Systems*, Bulgaria, 8-13
24. B. Kosko (1992) *Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence*, Prentice-Hall
25. M. Vaskovic (2014) *Compensation of nonlinear distortion in RF amplifiers for mobile communications*, PhD thesis, University of Westminster, United Kingdom
26. S.H. Melvin, M. Baro, M. Jandali, J. Ilow, (2008) Improved Compensation of HPA Nonlinearities using Digital Predistorters with Dynamic and Multi-Dimensional LUTs. *Proc. of the 6th Annual Communication Networks and Services Research Conference, CNSR 2008*, 46-52
27. T.K. Roy, M. Morshed, (2013) High Power Amplifier Effects Analysis for OFDM System. *International Journal of Science, Engineering and Technology Research (IJSETR)*, 2(5), 1119-1121
28. H. Zhi-yong, G. Jian-hua, G. Shu-jian, and W. Gang (2006) An improved look-up table predistortion technique for HPA with memory effects in OFDM systems. *IEEE Trans. Broadcast.* 52(1) : 87–91
29. M. Li, J. Liu, Y. Jiang, and W. Feng (2012) Complex-Chebyshev Functional Link Neural Network Behavioural Model for Broadband Wireless Power Amplifiers. *IEEE Transactions on Microwave Theory and Techniques*, 60(6) : 1979-1989
30. K. Fu, C.L. Law, T.T. Thein (2013) Novel Neural Network Model of Power Amplifier plus IQ Imbalances. *Progress in Electromagnetics Research B*, 46 : 177-192
31. V.S. Kodogiannis, E.M. Anagnostakis (2002) *Fuzzy Sets and Systems* 128(3) : 413-426
32. C-T. Lin, C.S. George Lee (1996) *Neural Fuzzy Systems: A Neuro-Fuzzy Synergism to Intelligent Systems*, Prentice Hall